Phase transition and percolation in Gibbsian particle models

Hans-Otto Georgii
Mathematisches Institut der Universität München
Theresienstr. 39, D-80333 München, Germany.

We discuss the interrelation between phase transitions in interacting lattice or continuum models, and the existence of infinite clusters in suitable random-graph models. In particular, we describe a random-geometric approach to the phase transition in the continuum Ising model of two species of particles with soft or hard interspecies repulsion. We comment also on the related area-interaction process and on perfect simulation.

1 Gibbs measures: general principles

This section contains a brief introduction to the basic physical and stochastic ideas leading to the concept of Gibbs measures. The principal question is the following:

> Which kind of stochastic model is appropriate for the description of spatial random phenomena involving a very large number of components which are coupled together by an interaction depending on their relative position?

To find an answer we will start with a spatially discrete situation; later we will proceed to the continuous case. Consider the phenomenon of ferromagnetism. A piece of ferromagnetic material like iron or nickel can be imagined as consisting of many elementary magnets, the so-called spins, which are located at the sites of a crystal lattice and have a finite number of possible orientations (according to the symmetries of the crystal). The essential point is that these spins interact with each other in such a way that neighboring spins prefer to be aligned. This interaction is responsible for the phenomenon of spontaneous magnetization, meaning that at sufficiently low temperatures the system can choose between several distinct macrostates in which typically all spins have the same orientation.

How can one find a mathematical model for such a ferromagnet? The first fact to observe is that the number of spins is very large. So, probabilistic experience with the law of large numbers suggests to approximate the large finite system by an infinite system in order to get clear-cut phenomena. This means that we should assume that the underlying crystal lattice is infinite. The simplest case to think of is the d-dimensional hypercubic lattice $\mathbf{Z}^d$. (As the case $d = 1$ is rather trivial, we will always assume that $d \geq 2$.) On the other hand, to keep the model simple it is natural to assume that each spin has only finitely many possible orientations. In other words, the random

spin ξ_i at lattice site i takes values in a finite state space S. The set of all possible spin configurations $\xi = (\xi_i)_{i\in\mathbf{Z}^d}$ is then the product space $\Omega = S^{\mathbf{Z}^d}$. This so-called *configuration space* is equipped with the Borel σ-algebra $\mathcal{F}$ for the natural product topology on Ω. Since the spins are random, we are interested in probability measures P on $(\Omega, \mathcal{F})$. Such probability spaces are known as *lattice systems*. For any $\xi \in \Omega$ and $\Lambda \subset \mathbf{Z}^d$ we write $\xi_\Lambda = (\xi_i)_{i\in\Lambda}$ for the part of the configuration that occurs in Λ. By abuse of notation, we use the same symbol ξ_Λ for the projection from Ω onto S^Λ.

Which kind of probability measure on $(\Omega, \mathcal{F})$ can serve as a model of a ferromagnet? As we have seen above, the essential feature of a ferromagnet is the interaction between the spins. We are thus interested in probability measures P on Ω for which the spin variables ξ_i, $i \in \mathbf{Z}^d$, are *dependent*. A natural way of describing dependencies is to prescribe certain conditional probabilities. This idea, which is familiar from Markov chains, turns out to be suitable also here. Since our parameter set $\mathbf{Z}^d$ admits no natural linear order, the conditional probabilities can, of course, not lead from a past to a future. Rather we *prescribe the behavior of a finite set of spins when all other spins are fixed*. In other words, we are interested in probability measures P on $(\Omega, \mathcal{F})$ having prescribed conditional probabilities

$$G_\Lambda(\xi_\Lambda|\xi_{\Lambda^c}) \tag{1}$$

for a configuration $\xi_\Lambda \in S^\Lambda$ within a finite set $\Lambda \subset \mathbf{Z}^d$ given a fixed configuration $\xi_{\Lambda^c} \in S^{\Lambda^c}$ off Λ. In the following we write $\Lambda \subset\subset \mathbf{Z}^d$ when Λ is a *finite* subset of $\mathbf{Z}^d$. The specific form of these conditional distributions does not matter at the moment. Two special cases are

- *the Markovian case:* the conditional distribution (1) only depends on the value of the spins along the *boundary* $\partial\Lambda = \{i \notin \Lambda : |i-j| = 1 \text{ for some } j \in \Lambda\}$ of Λ, i.e.,

  $$G_\Lambda(\xi_\Lambda|\xi_{\Lambda^c}) = G_\Lambda(\xi_\Lambda|\xi_{\partial\Lambda}) \tag{2}$$

  (with a slight abuse of notation); $|\cdot|$ stands for the Euclidean norm.

- *the Gibbsian case:* the conditional distribution (1) is defined in terms of a *Hamilton function* H_Λ by the Boltzmann–Gibbs formula

  $$G_\Lambda(\xi_\Lambda|\xi_{\Lambda^c}) = Z^{-1}_{\Lambda|\xi_{\Lambda^c}} \exp[-H_\Lambda(\xi)]\ , \tag{3}$$

  where $Z_{\Lambda|\xi_{\Lambda^c}} = \sum_{\xi'\in\Omega:\xi'\equiv\xi \text{ off } \Lambda} \exp[-H_\Lambda(\xi')]$ is a normalizing constant. Physically, $H_\Lambda(\xi)$ describes the energy excess of the total configuration ξ over the energy of the outer configuration ξ_{Λ^c}. (Physicists will miss here the factor β, the inverse temperature; we will assume that β is subsumed into H_Λ or, equivalently, that the units are chosen in such a way that $\beta = 1$.)

In the following, $G_\Lambda(\,\cdot\,|\xi_{\Lambda^c})$ will be viewed as a probability measure on Ω for which the configuration outside Λ is almost surely equal to ξ_{Λ^c}.

The above idea of prescribing conditional probabilities leads to the following concept introduced in the late 1960's independently by R.L. Dobrushin, and O.E. Lanford and D. Ruelle.

Definition 1.1 *A probability measure P on $(\Omega, \mathcal{F})$ is called a Gibbs measure, or DLR-state, for a family $\mathbf{G} = (G_\Lambda)_{\Lambda \subset\subset \mathbf{Z}^d}$ of conditional probabilities (satisfying the natural consistency condition) if*

$$P(\xi_\Lambda \text{ occurs in } \Lambda \,|\, \xi_{\Lambda^c} \text{ occurs off } \Lambda) = G_\Lambda(\xi_\Lambda \,|\, \xi_{\Lambda^c})$$

for P-almost all ξ_{Λ^c} and all $\Lambda \subset\subset \mathbf{Z}^d$.

If $\mathbf{G}$ is Gibbsian for a Hamiltonian H as in (3), each Gibbs measure can be interpreted as an equilibrium state for a physical system with state space S and interaction H. This is because the Boltzmann–Gibbs distribution maximizes the entropy when the mean energy is fixed; we will discuss this point in more detail in the continuum setting in Section 4.2 below.

A general account of the theory of Gibbs measures can be found in the monograph Georgii (1988); here we will only present the principal ideas. In contrast to the situation for Markov chains, Gibbs measures do not exist automatically. However, in the present case of a finite state space S, Gibbs measures do exist whenever $\mathbf{G}$ is Markovian in the sense of (2), or almost Markovian in the sense that the conditional probabilities (1) are continuous functions of the outer configuration ξ_{Λ^c}. In this case one can show that any weak limit of $G_\Lambda(\,\cdot\,|\xi_{\Lambda^c})$ for fixed $\xi \in \Omega$ as $\Lambda \uparrow \mathbf{Z}^d$ is a Gibbs measure.

The basic observation is that the Gibbs measures for a given consistent family $\mathbf{G}$ of conditional probabilities form a convex set $\mathcal{G}$. Therefore one is interested in its extremal points. These can be characterized as follows.

Theorem 1.2 *Let $\mathcal{T} = \bigcap \sigma(\xi_{\Lambda^c} : \Lambda \subset\subset \mathbf{Z}^d)$ the tail σ-algebra, i.e., the σ-algebra of all macroscopic events not depending on the values of any finite set of spins. Then the following statements hold:*

(a) *A Gibbs measure $P \in \mathcal{G}$ is extremal in $\mathcal{G}$ if and only if P is trivial on $\mathcal{T}$, i.e., if and only if any tail measurable real function is P-almost surely constant.*

(b) *Any two distinct extremal Gibbs measure are mutually singular on $\mathcal{T}$.*

(c) *Any non-extremal Gibbs measure is the barycenter of a unique probability weight on the set of extremal Gibbs measures.*

A proof can be found in Georgii (1988), Theorems (7.7) and (7.26). Statement (a) means that the extremal Gibbs measures are *macroscopically deterministic*: on the macroscopic level all randomness disappears, and an experimenter will get non-fluctuating measurements of macroscopic quantities like magnetization or energy per lattice site. Statement (b) asserts that *distinct extremal Gibbs measures show different macroscopic behavior.* So, they can be distinguished by looking at typical realizations of the spin configuration through macroscopic glasses. Finally, statement (c) implies that any realization which is typical for a non-extremal Gibbs measure is in fact typical for a suitable extremal Gibbs measure. In physical terms: *any configuration which can be seen in nature is governed by an extremal Gibbs measure*, and the non-extremal Gibbs measures can only be interpreted in a Bayesian way as measures describing the uncertainty of the experimenter. These observations lead us to the following definition.

Definition 1.3 *Any extremal Gibbs measure is called a phase of the corresponding physical system. If distinct phases exist, one says that a phase transition occurs.*

So, in terms of this definition the existence of phase transition is equivalent to the non-uniqueness of the Gibbs measure. In the light of the preceding theorem, this corresponds to a "macroscopic ambivalence" of the system's behavior. We should add that not all critical phenomena in nature can be described in this way: even when the Gibbs measure is unique it may occur that it changes its qualitative behavior when some parameters of the interaction are changed. However, we will not discuss these possibilities here and stick to the definition above for definiteness. In this contribution we will ask:

> What are the driving forces giving rise to a phase transition? Is there any mechanism relating microscopic and macroscopic behavior of spins?

As we will see, in a number of cases one can give the following geometric answer:

> One such mechanism is the formation of infinite clusters in suitably defined random graphs. Such infinite clusters serve as a link between the local and global behavior of spins, and make visible how the individual spins unite to form a specific collective behavior.

In the next section we will discuss two lattice models for which this answer is correct. In Section 4 we will show that a similar answer can also be given in a continuum set-up. A useful technical tool is the stochastic comparison of probability measures.

Suppose the state space S is a subset of $\mathbf{R}$ and thus linearly ordered. Then the configuration space Ω has a natural partial order, and we can speak of increasing real functions. Let P, P' be two probability measures on Ω. We say that P is *stochastically smaller* than P', and write $P \preceq P'$, if $\int f\, dP \le \int f\, dP'$ for all local increasing functions (or, equivalently, for all measurable bounded increasing functions) f on Ω. A sufficient condition for stochastic monotonicity is given in the proposition below. Although this condition refers to the case of finite products (for which stochastic monotonicity is similarly defined), it is also useful in the case of infinite product spaces. This is because (by the very definition) the relation $\preceq$ is preserved under weak limits.

Proposition 1.4 (Holley's inequality) *Let S be a finite subset of $\mathbf{R}$, Λ a finite index set, and P, P' two probability measures on the finite product space S^Λ giving positive weight to each element of S^Λ. Suppose the single-site conditional probabilities at any $i \in \Lambda$ satisfy*

$$P(\,\cdot\,|\xi_{\Lambda\setminus\{i\}} \text{ occurs off } i) \preceq P'(\,\cdot\,|\xi'_{\Lambda\setminus\{i\}} \text{ occurs off } i) \quad \textit{whenever } \xi \le \xi'.$$

Then $P \preceq P'$. If this condition holds with $P' = P$ then P has positive correlations in the sense that any two bounded increasing functions are positively correlated.

For a proof (and a slight extension) we refer to Theorems 4.8 and 4.11 of Georgii, Häggström and Maes (1999).

2 Phase transition and percolation: two lattice models

To provide the necessary background for our results on continuum particle systems let us still stick to the lattice case. We will consider two classical models which allow an understanding of phase transition in random-geometric terms. Many further examples for the relation between random geometry and phase transition can be found in Georgii, Häggström and Maes (1999).

Let us start recalling some basic facts on *Bernoulli percolation* on $\mathbf{Z}^d$ for $d \geq 2$. Consider $\mathbf{Z}^d$ as a graph with vertex set $\mathbf{Z}^d$ and edge set $E(\mathbf{Z}^d) = \{e = \{i,j\} \subset \mathbf{Z}^d : |i-j| = 1\}$. We fix two parameters $0 \leq p_s, p_b \leq 1$, the site and bond probabilities, and construct a random subgraph $\Gamma = (X, E)$ of $(\mathbf{Z}^d, E(\mathbf{Z}^d))$ by setting

$$X = \{i \in \mathbf{Z}^d : \xi_i = 1\}, \quad E = \{e \in E(X) : \eta_e = 1\},$$

where $E(X) = \{e \in E(\mathbf{Z}^d) : e \subset X\}$ is the set of all edges between the sites of X, and ξ_i, $i \in \mathbf{Z}^d$, and η_e, $e \in E(\mathbf{Z}^d)$, are independent Bernoulli variables satisfying $P(\xi_i = 1) = p_s$, $P(\eta_e = 1) = p_b$. This construction is called the *Bernoulli mixed site-bond percolation model*; setting $p_b = 1$ we obtain pure site percolation, and the case $p_s = 1$ corresponds to pure bond percolation.

Let $\{0 \leftrightarrow \infty\}$ denote the event that Γ contains an infinite path starting from 0, and

$$\theta(p_s, p_b; \mathbf{Z}^d) = \mathrm{Prob}(0 \leftrightarrow \infty)$$

be its probability. By Kolmogorov's zero-one law, we have $\theta(p_s, p_b; \mathbf{Z}^d) > 0$ if and only if Γ contains an infinite cluster with probability 1. In this case one says that percolation occurs. The following proposition asserts that this happens in a non-trivial region of the parameter square, which is separated by the so-called critical line from the region where all clusters of Γ are almost surely finite. The change of behavior at the critical line is the simplest example of a critical phenomenon.

Proposition 2.1 *The function $\theta(p_s, p_b; \mathbf{Z}^d)$ is increasing in p_s, p_b and d. Moreover, $\theta(p_s, p_b; \mathbf{Z}^d) = 0$ when $p_s p_b$ is small enough, while $\theta(p_s, p_b; \mathbf{Z}^d) > 0$ when $d \geq 2$ and $p_s p_b$ is sufficiently close to 1.*

Sketch proof: The monotonicity in p_s and p_b follows from Proposition 1.4, and the one in d from an obvious embedding argument. To show that $\theta = 0$ when $p_s p_b$ is small, we note that the expected number of neighbors in Γ of a given lattice site is $2dp_s p_b$. Comparison with a branching process thus shows that $\theta = 0$ when $2dp_s p_b < 1$.

Next, let $d = 2$ and suppose $0 \in X$ but the cluster C_0 of Γ containing 0 is finite. Consider $\partial_{ext} C_0$, the part of ∂C_0 belonging to the infinite component of C_0^c. For each site $i \in \partial_{ext} C_0$, either this site or all bonds leading from i to C_0 do not belong to Γ. This occurs with probability at most $1 - p_s p_b$. So, the probability that $\partial_{ext} C_0$ has a fixed location is at most $(1 - p_s p_b)^\ell$ with $\ell = \#\partial_{ext} C_0$. Counting all possibilities for this location one finds that $1 - \theta < 1$ when $1 - p_s p_b$ is small enough. By the monotonicity in d, the same holds a fortiori in higher dimensions. $\square$

The above proposition is all what we need here on Bernoulli percolation; an excellent source for a wealth of further rigorous results is the book of Grimmett (1999).

We now ask for the role of percolation for Gibbs measures, and in particular for the existence of phase transitions. Of course, in contrast to the above this will involve *dependent*, i.e., non-Bernoulli percolation. We consider here two specific examples. In these examples, the family **G** of conditional probabilities will be Gibbsian for a nearest-neighbor interaction; this means that both (2) and (3) are valid.

2.1 The Ising model

This is by far the most famous model of Statistical Mechanics, named after E. Ising who studied this model in the early 1920s in his thesis suggested by W. Lenz. It is the simplest model of a ferromagnet in equilibrium. One assumes that the spins have only two possible orientations, and therefore defines $S = \{-1, 1\}$. The family **G** is defined by (3) with

$$H_\Lambda(\xi) = J \sum_{\{i,j\}\cap\Lambda\neq\emptyset:|i-j|=1} 1_{\{\xi_i\neq\xi_j\}} \;, \tag{4}$$

where $J > 0$ is a coupling constant which is inversely proportional to the absolute temperature. This means that neighboring spins of different sign have to pay an energy cost J. There exist two configurations of minimal energy, so-called ground states, namely the configuration '$+$' which is identically equal to $+1$, and the configuration '$-$' identically equal to -1. The behavior of the model is governed by these two ground states. To see this we begin with some useful consequences of the ferromagnetic character of the interaction. First, it is intuitively obvious that the measures $G_\Lambda^+ = G_\Lambda(\,\cdot\,|+)$ decrease stochastically when Λ increases (since then the effect of the $+$ boundary decreases). This follows easily from Holley's inequality, Proposition 1.4. Since the local increasing functions are a convergence determining class, it follows that the weak infinite-volume limit $P^+ = \lim_{\Lambda\uparrow\mathbf{Z}^d} G_\Lambda^+$ exists. Likewise, the weak limit $P^- = \lim_{\Lambda\uparrow\mathbf{Z}^d} G_\Lambda(\,\cdot\,|-)$ exists (and by symmetry is the image of P^+ under simultaneous spin flip). These limits are Gibbs measures and invariant under translations. Holley's inequality also implies that they are stochastically maximal resp. minimal in $\mathcal{G}$, and in particular extremal. This gives us the following criterion for phase transition in the Ising model.

Proposition 2.2 *For the Ising model on $\mathbf{Z}^d$ with any coupling constant $J > 0$ we have $\#\mathcal{G} > 1$ if and only if $P^- \neq P^+$ if and only if $\int \xi_0 \, dP^+ > 0$.*

The last equivalence follows from the relation $P^- \preceq P^+$, the translation invariance of these Gibbs measures, and the spin-flip symmetry. A detailed proof of the proposition and the previous statements can be found in Section 4.3 of Georgii, Häggström and Maes (1999).

How can we use this criterion? This is where random geometry enters the scenery. The key is the following geometric construction tracing back to Fortuin and Kasteleyn (1972) and in this form to Edwards and Sokal (1988). It is called the *random-cluster representation of the Ising model.*

Let $\mathcal{E}_\Lambda^+ = \{E \subset E(\mathbf{Z}^d) : E \supset E(\Lambda^c)\}$ be the set of all edge configurations in $\mathbf{Z}^d$ which include all edges outside Λ, and define a probability measure ϕ_Λ on $\mathcal{E}_\Lambda^+$ by setting

$$\phi_\Lambda(E) = Z_\Lambda^{-1}\, 2^{k(E)}\, p^{\#E\setminus E(\Lambda^c)}\, (1-p)^{\#E(\mathbf{Z}^d)\setminus E} \quad \text{when } E \supset E(\Lambda^c), \tag{5}$$

where $p = 1 - e^{-J}$, $k(E)$ is the number of clusters of the graph $(\mathbf{Z}^d, E)$, and Z_Λ is a normalizing constant. ϕ_Λ is called the *random-cluster distribution in* Λ *with wired boundary condition.* This measure turns out to be related to $G_\Lambda^+ = G_\Lambda(\,\cdot\,|+)$. It will be convenient to identify a configuration $\xi \in \Omega$ with the pair (X^+, X^-), where X^+ and X^- are the sets of all lattice sites i for which $\xi_i = +1$ resp. -1.

Proposition 2.3 *For any hypercube* Λ *in* $\mathbf{Z}^d$ *there exists the following correspondence between the the Gibbs distribution* G_Λ^+ *for the Ising model and the random-cluster distribution* ϕ_Λ *in (5).*

$(G_\Lambda^+ \rightsquigarrow \phi_\Lambda)$ *Take a spin configuration* $\xi = (X^+, X^-) \in \Omega$ *with distribution* G_Λ^+*, and define an edge configuration* $E \in \mathcal{E}_\Lambda^+$ *as follows: Independently for all* $e \in E(\mathbf{Z}^d)$ *let* $e \in E$ *with probability*

$$p_\Lambda(e) = \begin{cases} 1 - e^{-J} & \text{if } e \subset X^+ \text{ or } e \subset X^-, \text{ and } e \cap \Lambda \neq \emptyset, \\ 1 & \text{if } e \subset \Lambda^c, \\ 0 & \text{otherwise,} \end{cases}$$

and $e \notin E$ *otherwise. Then* E *has distribution* ϕ_Λ*.*

$(\phi_\Lambda \rightsquigarrow G_\Lambda^+)$ *Pick an edge configuration* $E \in \mathcal{E}_\Lambda^+$ *according to* ϕ_Λ*, and define a spin configuration* $\xi = (X^+, X^-) \in \Omega$ *as follows: For each finite cluster* C *of* $(\mathbf{Z}^d, E)$ *let* $C \subset X^+$ *or* $C \subset X^-$ *according to independent flips of a fair coin; the unique infinite cluster of* $(\mathbf{Z}^d, E)$ *containing* Λ^c *is included into* X^+*. Then* ξ *has distribution* G_Λ^+*.*

Proof. A joint description of a spin configuration $\xi \in \Omega$ with distribution G_Λ^+ and an edge configuration $E \in \mathcal{E}_\Lambda^+$ with distribution ϕ_Λ can be obtained as follows. Any edge $e \in E(\mathbf{Z}^d)$ is independently included into E with probability $p = 1 - e^{-J}$ resp. 1 according to whether $e \cap \Lambda \neq \emptyset$ or not; each spin in Λ is equal to $+1$ or -1 according to independent flips of a fair coin; the spins off Λ are set equal to $+1$. The measure P thus described is then conditioned on the event A that no spins of different sign are connected by an edge. Relative to $P(\,\cdot\,|A)$, ξ has distribution G_Λ^+ and E has distribution ϕ_Λ. This is because $\exp[-H_\Lambda(\xi)]$ is equal to the conditional P-probability of A given ξ, and $2^{k(E)-1}$ is proportional to the conditional P-probability of A given E. Now, it is easy to see that the two constructions in the proposition simply correspond to the conditional distributions of E given ξ resp. of ξ given E relative to $P(\,\cdot\,|A)$. $\square$

Intuitively, the edges in the random-cluster representation indicate which pairs of spins "realize" their interaction, in that they decide to take the same orientation to avoid the dealignment costs. On the one hand, this representation is the basis of an efficient simulation procedure, the algorithm of Swendsen–Wang (1987), which together with its continuous counterpart will be discussed at the end of Section 4.4. On the other hand, it is the key for a geometric approach to the phenomenon of phase transition, as we will now show.

The construction $\phi_\Lambda \rightsquigarrow G_\Lambda^+$ implies that, for $0 \in \Lambda$, the conditional expectation of ξ_0 given E is 1 when 0 is connected to $\partial\Lambda$ by edges in E, and 0 otherwise. Hence

$$\int \xi_0 \, dP_\Lambda^+ = \phi_\Lambda(0 \leftrightarrow \partial\Lambda) \ . \tag{6}$$

By Proposition 1.4, the measures ϕ_Λ decrease stochastically when Λ increases, so that the infinite-volume random-cluster distribution $\phi = \lim_{\Lambda\uparrow\mathbf{Z}^d} \phi_\Lambda$ exists. Letting $\Lambda \uparrow \mathbf{Z}^d$ in (6) we thus find that $\int \xi_0 \, dP^+ = \phi(0 \leftrightarrow \infty)$. Combining this with Proposition 2.2 we obtain the first statement of the following theorem, the equivalence of percolation and phase transition. This gives us detailed information on the existence of phase transition.

Theorem 2.4 *Consider the Ising model on $\mathbf{Z}^d$ with Hamiltonian (4) for any coupling constant $J > 0$. Then $\#\mathcal{G} > 1$ if and only if $\phi(0 \leftrightarrow \infty) > 0$. Consequently, there exists a coupling threshold $0 < J_c < \infty$ (corresponding to a critical inverse temperature) such that $\#\mathcal{G} = 1$ when $J < J_c$ and $\#\mathcal{G} > 1$ when $J > J_c$.*

Sketch proof. It only remains to show the second statement. This follows from Holley's inequality, Proposition 1.4. First, this inequality implies that ϕ_Λ is stochastically increasing in the parameter $p = 1 - e^{-J}$. Hence $\phi(0 \leftrightarrow \infty)$ is an increasing function of p. Moreover, one finds that ϕ is stochastically dominated by the Bernoulli bond percolation measure, whence $\phi(0 \leftrightarrow \infty) = 0$ when p is so small that $\theta(1, p; \mathbf{Z}^d) = 0$. Finally, ϕ is stochastically larger than the Bernoulli bond percolation measure with parameter $\tilde{p} = p/(p + 2(1-p))$. Hence $\phi(0 \leftrightarrow \infty) > 0$ when p is so large that $\theta(1, \tilde{p}; \mathbf{Z}^d) > 0$. Details of this computation can be found in Section 6 of Georgii, Häggström and Maes (1999), which deals in fact with the extension of these results to the Potts model in which each spin has q different values. □

One may ask whether the connection between percolation and phase transition can be seen more directly in the behavior of spins. The following corollary gives an answer to this question. Let $\{0 \overset{+}{\longleftrightarrow} \infty\}$ denote the event that 0 belongs to an infinite cluster of the graph $(X^+, E(X^+))$ induced by the set of plus-spins.

Corollary 2.5 *For the Ising model on $\mathbf{Z}^d$ with arbitrary coupling constant $J > 0$ we have $P^+(0 \overset{+}{\longleftrightarrow} \infty) > 0$ whenever $\#\mathcal{G} > 1$. The converse holds only when $d = 2$.*

The first part follows readily from the construction in Proposition 2.3 which shows that $P^+(0 \overset{+}{\longleftrightarrow} \infty) \geq \phi(0 \leftrightarrow \infty)$. For its second part see Georgii, Häggström and Maes (1999). Pursuing the idea of plus-percolation further one can obtain the following result independently obtained in the late 1970s by Aizenman and Higuchi on the basis of previous work of L. Russo; a simpler proof has recently been given by Georgii and Higuchi (1999).

Theorem 2.6 *For the Ising model on $\mathbf{Z}^2$ with $J > J_c$, there exist no other phases than P^+ and P^-.*

A celebrated result of Dobrushin asserts that in three or more dimensions there exist non-translation invariant Gibbs measures which look like P^+ in one half-space and like P^- in the other half-space.

2.2 The Widom–Rowlinson lattice gas

The Widom–Rowlinson lattice gas is a discrete analog of a continuum model to be considered in Section 4. It describes the random configurations of particles of two

different types, plus or minus, which can only sit at the sites of the lattice $\mathbf{Z}^d$. Multiple occupations are excluded. So, at each site i of the lattice there are three possibilities: either i is occupied by a plus-particle, or by a minus-particle, or i is empty. The configuration space is thus $\Omega = S^{\mathbf{Z}^d}$ with $S = \{-1, 0, 1\}$. The basic assumption is that there is a hard-core repulsion between plus- and minus-particles, which means that particles of distinct type are not allowed to sit next to each other. In addition, there exists a chemical "activity" $z > 0$ which governs the overall-density of particles. The Hamiltonian thus takes the form

$$H_\Lambda(\xi) = \sum_{\{i,j\}\cap\Lambda\neq\emptyset : |i-j|=1} U(\xi_i, \xi_j) - \log z \sum_{i\in\Lambda} |\xi_i| \,, \tag{7}$$

where $U(\xi_i, \xi_j) = \infty$ if $\xi_i\xi_j = -1$, and $U(\xi_i, \xi_j) = 0$ otherwise. The associated family $\mathbf{G}$ of conditional probabilities is again given by (3). Just as in the Ising model, for $z > 1$ there exist two distinguished configurations of minimal energy, namely the constant configurations '$+$' and '$-$' for which all sites are occupied by particles of the same type. Moreover, one can again apply Holley's inequality to show that the Gibbs distributions $G_\Lambda^+ = G_\Lambda(\,\cdot\,|+)$ and $G_\Lambda^- = G_\Lambda(\,\cdot\,|-)$ converge to translation invariant limits $P^+, P^- \in \mathcal{G}$ which are stochastically maximal resp. minimal in $\mathcal{G}$, and therefore extremal. This implies that Proposition 2.2 holds verbatim also in the present case.

Is there also a geometric representation of the model, just as for the Ising model? The answer is yes, with interesting analogies and differences. There exists again a random-cluster distribution with an appearance very similar to (5), but this involves site percolation rather than bond percolation. Namely, consider the probability measure ψ_Λ on the set $\mathcal{X}_\Lambda^+ = \{Y \subset \mathbf{Z}^d : Y \supset \Lambda^c)\}$ which is given by

$$\psi_\Lambda(Y) = Z_\Lambda^{-1}\, 2^{k(Y)}\, p^{\#Y\setminus\Lambda^c}\, (1-p)^{\#\mathbf{Z}^d\setminus Y} \qquad \text{for } Y \supset \Lambda^c\,; \tag{8}$$

here $p = \frac{z}{1+z}$, $k(Y)$ is the number of clusters of the graph $(Y, E(Y))$, and Z_Λ is a normalizing constant. ψ_Λ is called the *site random-cluster distribution in* Λ *with parameter* p *and wired boundary condition*. We will again identify a configuration $\xi \in \Omega$ with a pair (X^+, X^-), where X^+ and X^- are the sets of all lattice points i such that $\xi_i = +1$ resp. -1; thus $X^+ \cup X^-$ is the set of occupied sites, and $\xi \equiv 0$ on its complement. Here is the *random-cluster representation of the Widom–Rowlinson lattice gas* which is analogous to Proposition 2.3.

Proposition 2.7 *For any hypercube* Λ *in* $\mathbf{Z}^d$ *there exists the following correspondence between the the Gibbs distribution* G_Λ^+ *for the Widom–Rowlinson model and the site random-cluster distribution* ψ_Λ *in (8).*

($G_\Lambda^+ \rightsquigarrow \psi_\Lambda$) *For a random spin configuration* $\xi = (X^+, X^-) \in \Omega$ *with distribution* G_Λ^+*, the random set* $Y = X^+ \cup X^-$ *has distribution* ψ_Λ*.*

($\psi_\Lambda \rightsquigarrow G_\Lambda^+$) *Pick a random set* $Y \in \mathcal{X}_\Lambda^+$ *according to* ψ_Λ*, and define a spin configuration* $\xi = (X^+, X^-) \in \Omega$ *with* $X^+ \cup X^- = Y$ *as follows: For each finite cluster* C *of* $(Y, E(Y))$ *let* $C \subset X^+$ *or* $C \subset X^-$ *according to independent flips of a fair coin; the unique infinite cluster of* $(Y, E(Y))$ *containing* Λ^c *is included into* X^+*. Then* ξ *has distribution* G_Λ^+*.*

The random-cluster representation of the Widom-Rowlinson model is simpler than that of the Ising model because the randomness involves only the sites of the lattice. (This is a consequence of the hard-core interaction; in the case of a soft repulsion the situation would be different, as we will see in the continuum setting in Section 4.) On the other hand, there is a serious drawback of the site random-cluster distribution ψ_Λ: it does not satisfy the conditions of Proposition 1.4 for positive correlations. This is because the conditional probabilities in (10) below are not increasing in Y. So, we still have a counterpart to (6), viz.

$$\int \xi_0 \, dG_\Lambda^+ = \psi_\Lambda(0 \leftrightarrow \partial\Lambda) \; , \tag{9}$$

but we do not know if the measures ψ_Λ are stochastically increasing in z and stochastically decreasing in Λ. So we obtain a somewhat weaker theorem.

Theorem 2.8 *Consider the Widom–Rowlinson model on $\mathbf{Z}^d$, $d \geq 2$, defined by (7) with activity $z > 0$. Then $\#\mathcal{G} > 1$ if and only if $\lim_{\Lambda\uparrow\mathbf{Z}^d} \psi_\Lambda(0 \leftrightarrow \partial\Lambda) > 0$. In particular, we have $\#\mathcal{G} = 1$ when z is sufficiently small, and $\#\mathcal{G} > 1$ when z is large enough.*

Sketch proof. The first statement follows immediately from (9) and the analog of Proposition 2.2. To prove the second statement we note that ψ_Λ has single-site conditional probabilities of the form

$$\psi_\Lambda(Y \ni i \,|\, Y \setminus \{i\}) = p \Big/ \big[p + (1-p)\, 2^{\kappa(i,Y)-1}\big] \; , \tag{10}$$

where $\kappa(i, Y)$ is the number of clusters of $Y \setminus \{i\}$ that intersect a neighbor of i. Since $0 \leq \kappa(i, Y) \leq 2d$, it follows from Proposition 1.4 that ψ_Λ is stochastically dominated by the site-Bernoulli measure with parameter $p^* = p/(p + (1-p)2^{-1})$, and dominates the site-Bernoulli measure with parameter $p_* = p/(p + (1-p)2^{2d-1})$. Combining this with Proposition 2.1 we thus find that $\#\mathcal{G} = 1$ when z is so small that $\theta(p^*, 1; \mathbf{Z}^d) = 0$, and $\#\mathcal{G} > 1$ when z is so large that $\theta(p_*, 1; \mathbf{Z}^d) > 0$. $\square$

Let us note that, in contrast to Theorem 2.4, the preceding result does not extend to the case when there are more than two different types of particles; this is related to the lack of stochastic monotonicity in this model. However, due to its simpler random-cluster representation the Widom–Rowlinson model has one advantage over the Ising model, in that it satisfies a much stronger counterpart to Corollary 2.5. In analogy to the notation there we write $\{0 \overset{+}{\longleftrightarrow} \infty\}$ for the event that the origin belongs to an infinite cluster of plus-particles.

Corollary 2.9 *For the Widom–Rowlinson lattice gas on $\mathbf{Z}^d$ for arbitrary dimension $d \geq 2$ and with any activity $z > 0$, we have $\#\mathcal{G} > 1$ if and only if $P^+(0 \overset{+}{\longleftrightarrow} \infty) > 0$.*

Sketch proof. The construction in Proposition 2.7 readily implies that

$$\psi_\Lambda(0 \leftrightarrow \partial\Lambda) = G_\Lambda^+(0 \overset{+}{\longleftrightarrow} \partial\Lambda) \; .$$

Combining this with (9) and letting $\Lambda \uparrow \mathbf{Z}^d$ one obtains the result. $\square$

The above equivalence of phase transition and percolation even holds when $\mathbf{Z}^d$ is replaced by an arbitrary graph.

As noticed before Theorem 2.8, we have no stochastic monotonicity in the activity z, and therefore no activity threshold for the existence of a phase transition. We are thus led to ask if, at least, the particle density is an increasing function of z. This can be deduced from general thermodynamic principles relying on convexity of thermodynamic functions rather than stochastic monotonicity. This will be described in Section 4.2 in the continuum set-up.

3 Continuum percolation

In the rest of this contribution we will show that quite a lot of the preceding results and techniques carry over to models of point particles in Euclidean space. In this section we deal with a simple model of continuum percolation. Roughly speaking, this model consists of Poisson points which are connected by Bernoulli edges. To be precise, let $\mathcal{X}$ denote the set of all locally finite subsets X of $\mathbf{R}^d$. $\mathcal{X}$ is the set of all point configurations in $\mathbf{R}^d$, and is equipped with the usual σ-algebra generated by the counting variables $X \to \#X_\Lambda$ for bounded Borel sets $\Lambda \subset \mathbf{R}^d$; here we use the abbreviation $X_\Lambda = X \cap \Lambda$. Next, let $\mathcal{E}$ be the set of all locally finite subsets of $E(\mathbf{R}^d) = \{\{x,y\} \subset \mathbf{R}^d : x \neq y\}$. $\mathcal{E}$ is the set of all possible edge configurations and is equipped with an analogous σ-algebra. For $X \in \mathcal{X}$ let $E(X) = \{e \in E(\mathbf{R}^d) : e \subset X\}$ the set of all possible edges between the points of X, and $\mathcal{E}_X = \{E \in \mathcal{E} : E \subset E(X)\}$ the set of edge configurations between the points of X. We construct a random graph $\Gamma = (X, E)$ in $\mathbf{R}^d$ as follows.

- Pick a random point configuration $X \in \mathcal{X}$ according to the Poisson point process π^z on $\mathbf{R}^d$ with intensity $z > 0$.

- For given $X \in \mathcal{X}$, pick a random edge configuration $E \in \mathcal{E}_X$ according to the Bernoulli measure μ_X^p on $\mathcal{E}_X$ for which the events $\{E \ni e\}$, $e = \{x,y\} \in E(X)$, are independent with probability $\mu_X^p(E \ni e) = p(x-y)$; here $p : \mathbf{R}^d \to [0,1]$ is a given even measurable function.

The distribution of our random graph Γ is thus determined by the probability measure

$$P^{z,p}(dX, dE) = \pi^z(dX)\, \mu_X^p(dE) \tag{11}$$

on $\mathcal{X} \times \mathcal{E}$. It is called the *Poisson random-edge model*, or *Poisson random-connection model*, and has been introduced and studied first by M. Penrose (1991); a detailed account of its properties is given in Meester and Roy (1996).

A special case of particular interest is when $p(x-y)$ is equal to 1 when $|x-y| \leq 2r$ for some $r > 0$, and 0 otherwise. This means that any two points x and y are connected by an edge if and only if the balls $B_r(x)$ and $B_r(x)$ with radius r and center x resp. y overlap. The connectivity properties of the corresponding Poisson random-edge model are thus the same as those of the random set $\Xi = \bigcup_{x \in X} B_r(x)$, for random X with distribution π^z. This special case is therefore called the *Boolean model*, or the *Poisson blob model*.

Returning to the general case, we consider the *percolation probability* of a typical point. Writing $x \leftrightarrow \infty$ when x belongs to an infinite cluster of $\Gamma = (X, E)$, this is given

by the expression

$$\theta(z,p;\mathbf{R}^d) = \int \frac{\#\{x \in X_\Lambda : x \leftrightarrow \infty\}}{|\Lambda|}\, P^{z,p}(dX,dE) \tag{12}$$

for an arbitrary bounded box Λ with volume $|\Lambda|$. By translation invariance, $\theta(z,p;\mathbf{R}^d)$ does not depend on Λ. In fact, in terms of the Palm measure $\hat{\pi}^z$ of π^z and the associated measure $\hat{P}^{z,p}(dX,dE) = \hat{\pi}^z(dX)\,\mu_X(dE)$ we can write

$$\theta(z,p;\mathbf{R}^d) = \hat{P}^{z,p}(0 \leftrightarrow \infty)\ .$$

The following result of M. Penrose (1991) is the continuum analog of Proposition 2.1.

Theorem 3.1 *$\theta(z,p;\mathbf{R}^d)$ is an increasing function of the intensity z and the edge probability function $p(\cdot)$. Moreover, $\theta(z,p;\mathbf{R}^d) = 0$ when $z\int p(x)\,dx$ is sufficiently small, while $\theta(z,p;\mathbf{R}^d) > 0$ when $z\int p(x)\,dx$ is large enough.*

Sketch proof: The monotonicity follows from an obvious stochastic comparison argument. Since $z\int p(x)\,dx$ is the expected number of edges emanating from a given point, a branching argument shows that $\theta(z,p;\mathbf{R}^d) = 0$ when $z\int p(x)\,dx < 1$. It remains to show that $\theta(z,p;\mathbf{R}^d) > 0$ when $z\int p(x)\,dx$ is large enough. By scaling we can assume that $\int p(x)\,dx = 1$. For simplicity we will in fact suppose that p is bounded away from 0 in a neighborhood of the origin, i.e., $p(x-y) \geq \delta > 0$ whenever $|x-y| \leq 2r$. (The following is a special case of an argument of Georgii and Häggström (1996).)

We divide the space $\mathbf{R}^d$ into cubic cells $\Delta(i)$, $i \in \mathbf{Z}^d$, with diameter at most r. We also pick a sufficiently large number n and introduce the following two concepts.

- Call a cell $\Delta(i)$ *good* if it contains at least n points which form a connected set relative to the edges of Γ in between them. This event does not depend on the configurations in all other cells and has probability at least

  $$\pi^z(N_i \geq n)\ [1-(n-1)(1-\delta^2)^{n-2}] \equiv p_s\ ;$$

  here, N_i is the random number of points in cell $\Delta(i)$, and the second term in the square bracket is an estimate for the probability that one of the n points is not connected to the first point by a sequence of two edges. The essential fact is that p_s is arbitrarily close to 1 when n and z are large enough.

- Call two adjacent cells $\Delta(i), \Delta(j)$ *linked* if there exists an edge from some point in $\Delta(i)$ to some point in $\Delta(j)$. Conditionally on the event that $\Delta(i)$ and $\Delta(j)$ are good, this has probability at least $1-(1-\delta)^{n^2} = p_b$, which is also close to 1 when n is large enough.

Now the point is the following: whenever there exists an infinite cluster of linked good cells (i.e., an infinite cluster in the countable graph with vertices at the good cells and with edges between pairs of linked cells) then there exists an infinite cluster in the original Poisson random-edge model. Hence $\theta(z,p;\mathbf{R}^d) \geq \frac{n}{v}\,\theta(p_s,p_b;\mathbf{Z}^d)$, where v is the cell volume and $\theta(p_s,p_b;\mathbf{Z}^d)$ is as in Proposition 2.1. Hence $\theta(z,p;\mathbf{R}^d) > 0$ when z is large enough. $\square$

How can one extend a percolation result as above from the Poisson case to point processes with spatial dependencies? Just as in the lattice gas, one can take advantage of stochastic comparison techniques. To this end we need a continuum analog of Holley's theorem.

A simple point process P on a bounded Borel subset Λ of $\mathbf{R}^d$ (i.e., a probability measure on $\mathcal{X}_\Lambda = \{X \in \mathcal{X} : X \subset \Lambda\}$) is said to have Papangelou (conditional) intensity $\gamma : \Lambda \times \mathcal{X}_\Lambda \to [0, \infty[$ if P satisfies the identity

$$\int P(dX) \sum_{x \in X} f(x, X \setminus \{x\}) = \int dx \int P(dX)\, \gamma(x|X)\, f(x, X) \tag{13}$$

for any measurable function $f : \Lambda \times \mathcal{X}_\Lambda \to [0, \infty[$. (This is a non-stationary analog of the Georgii–Nguyen–Zessin equality discussed in the contribution of D. Stoyan to this volume.) This equation roughly means that $\gamma(x|X)\, dx$ is proportional to the conditional probability for the existence of a particle in an infinitesimal volume dx when the remaining configuration is X. Formally, it is not difficult to see that (13) is equivalent to the statement that P is absolutely continuous with respect to the intensity-1 Poisson point process $\pi_\Lambda = \pi_\Lambda^1$ in Λ with Radon–Nikodym density g satisfying $g(X \cup \{x\}) = \gamma(x|X)\, g(X)$, see e.g. Georgii and Küneth (1997). In particular, the Poisson process π_Λ^z of intensity $z > 0$ on Λ has Papangelou intensity $\gamma(x|X) = z$.

Proposition 3.2 (Holley-Preston inequality) *Let $\Lambda \subset \mathbf{R}^d$ be a bounded Borel set and P, P' two probability measures on $\mathcal{X}_\Lambda$ with Papangelou intensities γ resp. γ'. Suppose $\gamma(x|X) \le \gamma'(x|X')$ whenever $X \subset X'$ and $x \notin X' \setminus X$. Then $P \preceq P'$. If this condition holds with $P' = P$ then P has positive correlations.*

Of course, the stochastic partial order $P \preceq P'$ is defined by means of the inclusion relation on $\mathcal{X}_\Lambda$. Under additional technical assumptions the preceding proposition was first derived by Preston in 1975; in the present form it is due to Georgii and Küneth (1997). In the next section we will see how this result can be used to establish percolation in certain continuum random-cluster models, and thereby the existence of phase transitions in certain continuum particle systems.

4 The continuum Ising model

The continuum Ising model is a model of point particles in $\mathbf{R}^d$ of two different types, plus and minus. Rather than of particles of different types, one may also think of particles with a ferromagnetic spin with two possible orientations. The latter would be suitable for modelling ferrofluids such as the Au-Co alloy, which have recently found some physical attention. Much of what follows can also be extended to systems with more than two types, but we stick here to the simplest case. A configuration of particles is then described by a pair $\xi = (X^+, X^-)$, where X^+ and X^- are the configurations of plus- resp. minus-particles. The configuration space is thus $\Omega = \mathcal{X}^2$.

We assume that the particles interact via a repulsive interspecies pair potential of finite range, which is given by an even measurable function $J : \mathbf{R}^d \to [0, \infty]$ of

bounded support. The Hamiltonian in a bounded Borel set $\Lambda \subset \mathbf{R}^d$ of a configuration $\xi = (X^+, X^-)$ is thus given by

$$H_\Lambda(\xi) = \sum_{x \in X^+, y \in X^- :\, \{x,y\} \cap \Lambda \neq \emptyset} J(x-y) \ . \tag{14}$$

In view of its analogy to (4) this model is called the *continuum Ising model.* Setting $J(x-y) = \infty$ when $|x-y| \le 2r$ and $J(x-y) = 0$ otherwise, we obtain the classical *Widom–Rowlinson model* (1970) with a hard-core interspecies repulsion (which in spatial statistics is occasionally referred to as the *penetrable spheres mixture model*). Of course, this case corresponds to the Widom–Rowlinson lattice gas considered above. Here we make only the much weaker assumption that J is bounded away from zero on a neighborhood of the origin. That is, there exist constants $\delta, r > 0$ such that

$$J(x-y) \ge \delta \text{ when } |x-y| \le 2r. \tag{15}$$

An interesting generalization of the Hamiltonian (14) can be obtained by adding an interaction term which is independent of the types of the particles. In a ferrofluid model this would mean that in addition to the ferromagnetic interaction of particle spins there is also a molecular interaction which is spin-independent. Such an extension is considered in Georgii and Häggström (1996).

Given the Hamiltonian (14), the associated *Gibbs distribution in* Λ *with activity* $z > 0$ *and boundary condition* $\xi_{\Lambda^c} = (X^+_{\Lambda^c}, X^-_{\Lambda^c}) \in \mathcal{X}^2_{\Lambda^c}$ is defined by the formula

$$G_\Lambda(d\xi_\Lambda | \xi_{\Lambda^c}) = Z^{-1}_{\Lambda|\xi_{\Lambda^c}} \exp[-H_\Lambda(\xi)]\, \pi^z_\Lambda(dX^+_\Lambda)\, \pi^z_\Lambda(dX^-_\Lambda) \tag{16}$$

which is completely analogous to (3). The corresponding set $\mathcal{G} = \mathcal{G}(z)$ of Gibbs measures is then defined as in Definition 1.1 (with Λ running through the bounded Borel sets in $\mathbf{R}^d$ instead of the finite subsets of $\mathbf{Z}^d$).

In general, the existence of Gibbs measures in continuum models is not easy to establish. In the present case, however, it is simple: Thinking of $\mathcal{X}^2_\Lambda$ as the space of configurations on two disjoint copies Λ^+ and Λ^- of Λ, we see that $G_\Lambda(\,\cdot\,|\xi_{\Lambda^c})$ has the Papangelou intensity

$$\gamma(x|X^+, X^-) = \left\{ \begin{array}{ll} z \exp[-\sum_{y \in X^-} J(x-y)] & \text{if } x \in \Lambda^+ \\ z \exp[-\sum_{y \in X^+} J(x-y)] & \text{if } x \in \Lambda^- \end{array} \right\} \le z \ . \tag{17}$$

Proposition 3.2 therefore implies that $G_\Lambda(\,\cdot\,|\xi_{\Lambda^c}) \preceq \pi^z_\Lambda \times \pi^z_\Lambda$. Standard compactness theorems for point processes now show that for each $\xi \in \mathcal{X}^2$ the sequence $G_\Lambda(\,\cdot\,|\xi_{\Lambda^c})$ has an accumulation point P as $\Lambda \uparrow \mathbf{R}^d$, and it is easy to see that $P \in \mathcal{G}$.

4.1 Uniqueness and phase transition

We will now show that the Gibbs measure is unique when z is small, whereas a phase transition occurs when z is large enough. Both results rely on percolation techniques. As in the Widom–Rowlinson lattice gas, it remains open whether there is a sharp activity threshold separating intervals of uniqueness and non-uniqueness.

Proposition 4.1 *For the continuum Ising model we have $\#\mathcal{G}(z) = 1$ when z is sufficiently small.*

Sketch proof: Let $P, P' \in \mathcal{G}(z)$. We show that $P = P'$ when z is small enough. Let R be the range of J, i.e., $J(x) = 0$ when $|x| \leq R$, and divide $\mathbf{R}^d$ into cubic cells $\Delta(i)$, $i \in \mathbf{Z}^d$, of linear size R. Let p_c^* be the Bernoulli site percolation threshold of the graph with vertex set $\mathbf{Z}^d$ and edges between all points having distance 1 *in the max-norm.* Consider the Poisson measure $Q^z = \pi^z \times \pi^z$ on the configuration space $\Omega = \mathcal{X}^2$.

Let ξ, ξ' be two independent realizations of Q^z, and suppose z is so small that $Q^z \times Q^z(N_i + N_i' \geq 1) < p_c^*$, where N_i and N_i' are the numbers of particles (plus or minus) in ξ resp. ξ'. Then for any finite union Λ of cells we have $Q^z \times Q^z(\Lambda \overset{\geq 1}{\longleftrightarrow} \infty) = 0$, where $\{\Lambda \overset{\geq 1}{\longleftrightarrow} \infty\}$ denotes the event that a cell in Λ belongs to an infinite connected set of cells $\Delta(i)$ containing at least one particle in either ξ or ξ'. Proposition 3.2 together with (17) imply that $P \times P' \preceq Q^z \times Q^z$. Hence $P \times P'(\Lambda \overset{\geq 1}{\longleftrightarrow} \infty) = 0$. In other words, given two independent realizations ξ and ξ' of P and P' there exists a random corridor of width R around Λ which is completely free of particles. In particular, this means that ξ and ξ' coincide on this random corridor. By a spatial strong Markov property of Gibbs measures, it follows that P and P' coincide on the σ-algebra of events in Λ. As Λ can be chosen arbitrarily large, this proves the proposition. □

After this result on the absence of phase transition (following from the absence of some kind of percolation) we turn to the existence of phase transition. This will follow from the existence of percolation in a suitable random-cluster model. In analogy to Propositions 2.3 and 2.7, we will derive a random-cluster representation of the Gibbs distribution

$$G_\Lambda^+ = \int \pi_{\Lambda^c}^z(dY_{\Lambda^c}^+)\, G_\Lambda(\,\cdot\,|Y_{\Lambda^c}^+, \emptyset) \tag{18}$$

with a Poisson boundary condition of plus-particles and no minus-particle off Λ. Its random-cluster counterpart is the following probability measure χ_Λ on $\mathcal{X} \times \mathcal{E}$ describing random graphs (Y, E) in $\mathbf{R}^d$:

$$\chi_\Lambda(dY, dE) = Z_{\Lambda|Y_{\Lambda^c}}^{-1}\, 2^{k(Y,E)}\, \pi^z(dY)\, \mu_Y^{p,\Lambda}(dE)\ . \tag{19}$$

In the above, $k(Y, E)$ is the number of clusters of the graph (Y, E), $Z_{\Lambda|Y_{\Lambda^c}} = \int 2^{k(Y,E)}\, \pi_\Lambda^z(dY_\Lambda)$ normalizes the conditional probability of χ_Λ given $Y_{\Lambda^c} = Y \cap \Lambda^c$ (so that Y_{Λ^c} still has the Poisson distribution $\pi_{\Lambda^c}^z$), and $\mu_Y^{p,\Lambda}$ is the probability measure on $\mathcal{E}$ for which the edges $e = \{x, y\} \subset Y$ are drawn independently with probability $p(x - y) = 1 - e^{-J(x-y)}$ if $e \not\subset Y_{\Lambda^c}$, and probability 1 otherwise. The probability measure χ_Λ in (19) is called the *continuum random-cluster distribution in* Λ *with connection probability function* p *and wired boundary condition.* Note that (in contrast to (5) and (8)) this distribution describes random configurations of both points and edges. In the Widom–Rowlinson case of a hard-core interspecies repulsion the randomness of the edges disappears, and χ_Λ describes a dependent Boolean percolation model which is the direct continuum analog of (8). The *random-cluster representation of the continuum Ising model* now reads as follows.

Proposition 4.2 *For any bounded box Λ in $\mathbf{R}^d$ there is the following correspondence between the Gibbs distribution G_Λ^+ in (18) for the continuum Ising model and the random-cluster distribution χ_Λ in (19).*

($G_\Lambda^+ \rightsquigarrow \chi_\Lambda$) Take a particle configuration $\xi = (X^+, X^-) \in \Omega$ with distribution G_Λ^+ and define a random graph $(Y, E) \in \mathcal{X} \times \mathcal{E}$ as follows: Let $Y = X^+ \cup X^-$, and independently for all $e = \{x, y\} \in E(Y)$ let $e \in E$ with probability

$$p_\Lambda(e) = \begin{cases} 1 - e^{-J(x-y)} & \text{if } e \subset X^+ \text{ or } e \subset X^-, \text{ and } e \cap \Lambda \neq \emptyset, \\ 1 & \text{if } e \subset \Lambda^c, \\ 0 & \text{otherwise.} \end{cases}$$

Then (Y, E) has distribution χ_Λ.

($\chi_\Lambda \rightsquigarrow G_\Lambda^+$) Pick a random graph $(Y, E) \in \mathcal{X} \times \mathcal{E}$ according to χ_Λ. Define a particle configuration $\xi = (X^+, X^-) \in \Omega$ with $X^+ \cup X^- = Y$ as follows: For each finite cluster C of (Y, E) let $C \subset X^+$ or $C \subset X^-$ according to independent flips of a fair coin; the unique infinite cluster of (Y, E) containing Y_{Λ^c} is included into X^+. Then ξ has distribution G_Λ^+.

Just as in the lattice case, the random-cluster representation above gives the following key identity: for any finite box $\Delta \subset \Lambda$,

$$\begin{aligned} &\int [\#X_\Delta^+ - \#X_\Delta^-] \, G_\Lambda^+(dX^+, dX^-) \\ &\qquad = \int \#\{x \in Y_\Delta : x \leftrightarrow Y_{\Lambda^c}\} \, \chi_\Lambda(dY, dE) \; ; \end{aligned} \tag{20}$$

in the above, the notation $x \leftrightarrow Y_{\Lambda^c}$ means that x is connected to a point of Y_{Λ^c} in the graph (Y, E). In other words, the difference between the mean number of plus- and minus-particles in Δ corresponds to the percolation probability in χ_Λ. How can one check that the latter is positive for large z? The idea is again a stochastic comparison.

Let $\nu_\Lambda = \chi_\Lambda(\cdot \times \mathcal{E})$ the point marginal of χ_Λ. Then $\chi_\Lambda(dY, dE) = \nu_\Lambda(dY) \, \phi_{\Lambda,Y}(dE)$ with an obvious analog $\phi_{\Lambda,Y}$ of (5). An application of Proposition 1.4 shows that $\phi_{\Lambda,Y}$ is stochastically larger than the Bernoulli edge measure $\mu_Y^{\tilde{p}}$ for which edges are drawn independently between points $x, y \in Y$ with probability $\tilde{p}(x - y) = (1 - e^{-\delta})/[(1 - e^{-\delta}) + 2\, e^{-\delta}]$ when $|x - y| \le 2r$, and with probability 0 otherwise. Here, δ and r are as in assumption (15). Moreover, ν_Λ has the Papangelou intensity

$$\gamma(x|Y) = z \int 2^{k(Y \cup \{x\}, \, \cdot \,)} \, d\phi_{\Lambda, Y \cup \{x\}} \Big/ \int 2^{k(Y, \, \cdot \,)} \, d\phi_{\Lambda,Y} \; .$$

To get a lower estimate for $\gamma(x|Y)$ one has to compare the effect on the number of clusters in (Y, E) when a particle at x and corresponding edges are added. In principle, this procedure could connect a large number of distinct clusters lying close to x, so that $k(Y \cup \{x\}, \cdot\,)$ was much smaller than $k(Y, \cdot\,)$. However, one can show that this occurs only with small probability, so that $\gamma(x|Y) \ge \alpha z$ for some $\alpha > 0$. By Proposition 3.2, we can conclude that χ_Λ is stochastically larger than the Poisson random-edge measure $P^{\alpha z, \tilde{p}}$ defined in (11). The right-hand side of (20) is therefore not smaller than

$\theta(\alpha z, \tilde{p}; \mathbf{R}^d)$. Finally, since $G_\Lambda^+ \preceq \pi^z \times \pi^z$ by (17), the Gibbs distributions G_Λ^+ have a cluster point $P^+ \in \mathcal{G}(z)$ satisfying

$$\int [\# X_\Delta^+ - \# X_\Delta^-] \, P^+(dX^+, dX^-) \geq \theta(\alpha z, \tilde{p}; \mathbf{R}^d) \, .$$

By spatial averaging one can achieve that P^+ is in addition translation invariant. Together with Theorem 3.1 this leads to the following theorem.

Theorem 4.3 *For the continuum Ising model on $\mathbf{R}^d$, $d \geq 2$, with Hamiltonian (14) and sufficiently large activity z there exist two translation invariant Gibbs measures P^+ and P^- having a majority of plus- resp. minus-particles and related to each other by the plus-minus interchange.*

This result is due to Georgii and Häggström (1996). In the special case of the Widom-Rowlinson model it has been derived independently in the same way by Chayes, Chayes, and Kotecký (1995). The first proof of phase transition in the Widom-Rowlinson model was found by Ruelle in 1971, and for a soft but strong repulsion by Lebowitz and Lieb in 1972. Gruber and Griffiths (1986) used a direct comparison with the lattice Ising model in the case of a species-independent background hard core.

As a matter of fact, one can make further use of stochastic monotonicity. (In contrast to the preceding theorem, this only works in the present case of two particle types.) Introduce a partial order '$\leq$' on $\Omega = \mathcal{X}^2$ by writing

$$(X^+, X^-) \leq (Y^+, Y^-) \text{ when } X^+ \subset Y^+ \text{ and } X^- \supset Y^- . \tag{21}$$

A straightforward extension of Proposition 3.2 then shows that the measures G_Λ^+ in (18) decrease stochastically relative to this order when Λ increases. (This can be also deduced from the couplings obtained by perfect simulation, see Section 4.4 below.) It follows that P^+ is in fact the limit of these measures, and is in particular translation invariant. Moreover, one can see that P^+ is stochastically maximal in $\mathcal{G}$ in this order. This gives us the following counterpart to Corollary 2.9.

Corollary 4.4 *For the continuum Ising model with any activity $z > 0$, a phase transition occurs if and only if*

$$\int \hat{P}^+(dX^+, dX^-) \, \mu^p_{X^+}(0 \overset{+}{\longleftrightarrow} \infty) > 0 \, ;$$

here $\hat{P}^+$ is the Palm measure of P^+, and the relation $0 \overset{+}{\longleftrightarrow} \infty$ means that the origin belongs to an infinite cluster in the graph with vertex set X^+ and random edges drawn according to the probability function $p = 1 - e^{-J}$.

It is not known whether P^+ and P^- are the only extremal elements of $\mathcal{G}(z)$ when $d = 2$, as it is the case in the lattice Ising model. However, using a technique known in physics as the Mermin–Wagner theorem one can show the following.

Theorem 4.5 *If J is twice continuously differentiable then each $P \in \mathcal{G}(z)$ is translation invariant.*

A proof can be found in Georgii (1999). The existence of non-translation invariant Gibbs measures in dimensions $d \geq 3$ is an open problem.

4.2 Thermodynamic aspects

Although we were able to take some advantage of stochastic comparison techniques in the continuum Ising model, the use of Proposition 3.2 is much more limited than that of its lattice analog. The reason is that its condition requires some kind of attractivity, which is in conflict with stability (preventing the existence of infinitely many particles in a bounded region). This implies that a continuum Gibbs distribution G_Λ with a pair interaction cannot satisfy the conditions of Theorem 3.2 with $P = P' = G_\Lambda$, which would imply that G_Λ has positive correlations and is stochastically increasing with the activity z. Fortunately, this gap can be closed to some extent by the use of classical convexity techniques of Statistical Mechanics. These will allow us to conclude that at least the particle density of Gibbs measures is an increasing function of the activity z.

It should be noted that these ideas are standard in Statistical Physics; they are included here because they might be less known among spatial statisticians, and because we need to check that the general principles really work in the model at hand. One should also note that this technique does not depend on the specific features of the model; in particular, it applies also to the gas of hard balls discussed in H. Löwen's contribution to this volume.

Let us begin recalling the thermodynamic justification of Gibbs measures. Let $\Lambda \subset \mathbf{R}^d$ be a finite box. For any translation invariant probability measure P on $\Omega = \mathcal{X}^2$ consider the *entropy per volume*

$$\mathsf{s}(P) = \lim_{|\Lambda|\to\infty} |\Lambda|^{-1}\, S(P_\Lambda)\ .$$

Here we write P_Λ for the restriction of P to $\mathcal{X}^2_\Lambda$, the set of particle configurations in Λ, and

$$S(P_\Lambda) = \begin{cases} -\int \log f\ dP_\Lambda & \text{if } P_\Lambda \ll \pi^1_\Lambda \times \pi^1_\Lambda \text{ with density } f, \\ -\infty & \text{otherwise} \end{cases}$$

for the entropy of P_Λ relative two the two-species Poisson process $\pi^1_\Lambda \times \pi^1_\Lambda$ on $\mathcal{X}^2_\Lambda$ with intensity 1. The notation $|\Lambda| \to \infty$ means that Λ runs through a specified increasing sequence of cubic boxes with integer sidelength. The existence of $\mathsf{s}(P)$ is a multidimensional version of Shannon's theorem; see Georgii (1988) for the lattice case to which the present case can be reduced by identifying Ω with $(\mathcal{X}^2_C)^{\mathbf{Z}^d}$ for a unit cube C.

Next consider the *interaction energy per volume*

$$\mathsf{u}(P) = \int \hat{P}(dX^+, dX^-)\Big[1_{\{0\in X^+\}} \sum_{x\in X^-} J(x) + 1_{\{0\in X^-\}} \sum_{x\in X^+} J(x)\Big] \tag{22}$$

defined in terms of the Palm measure $\hat{P}$ of P. $\mathsf{u}(P)$ can also be defined as a per-volume limit, cf. Georgii (1994), Section 3. Also, consider the *particle density*

$$\varrho(P) = \hat{P}(\Omega) = |\Lambda|^{-1} \int [\#X^+_\Lambda + \#X^-_\Lambda]\, dP_\Lambda$$

of P; by translation invariance the last term does not depend on Λ. The term $-\varrho(P) \log z$ is then equal to the chemical energy per volume.

Finally, consider the *pressure*

$$\mathsf{p}(z) = -\min_P \Big[\mathsf{u}(P) - \varrho(P)\,\log z - \mathsf{s}(P)\Big]\ ; \tag{23}$$

the minimum extends over all translation invariant probability measures P on Ω. The large deviation techniques of Georgii (1994) show that $\mathsf{p}(z) = \lim_{|\Lambda|\to\infty} |\Lambda|^{-1} \log Z_{\Lambda|\xi_{\Lambda^c}}$ for each $\xi \in \Omega$. (The paper Georgii (1994) deals only with particles of a single type and superstable interaction, but the extension to the present case is straightforward because J is nonnegative and has finite range.) The variational principle for Gibbs measures then reads as follows.

Theorem 4.6 *Let P be a translation invariant probability measure on $\Omega = \mathcal{X}^2$. Then $P \in \mathcal{G}(z)$ if and only if $\mathsf{u}(P) - \varrho(P)\,\log z - \mathsf{s}(P)$, the free energy per volume, is equal to its minimum $-\mathsf{p}(z)$.*

The "only if" part can be derived along the lines of Georgii (1994) and Proposition 7.7 of Georgii (1995). The "if" part follows from the analogous lattice result (see Section 15.4 of Georgii (1988)) by the identification of Ω and $(\mathcal{X}_C^2)^{\mathbf{Z}^d}$ mentioned above.

What does the theorem tell us about the particle densities of Gibbs measures? Let us look at the pressure $\mathsf{p}(z)$. First, it follows straight from the definition (23) that $\mathsf{p}(z)$ is a convex function of $\log z$. In other words, the function $\tilde{\mathsf{p}}(t) = \mathsf{p}(e^t)$ is convex. Next, inserting $P = \pi^z \times \delta_\emptyset$ into the right-hand side of (23) we see that $\tilde{\mathsf{p}} > -\infty$, and that the slope of $\tilde{\mathsf{p}}$ at t tends to infinity as $t \to \infty$.

Now, suppose $P \in \mathcal{G}(z)$ is translation invariant. The variational principle above then implies that the function $t \to (t - \log z)\rho(P) + \mathsf{p}(z)$ is a tangent to $\tilde{\mathsf{p}}$ at $\log z$. For, on the one hand we have

$$\mathsf{u}(P) - \varrho(P)\,\log z - \mathsf{s}(P) = -\mathsf{p}(z) < \infty$$

and thus $\mathsf{u}(P) - \mathsf{s}(P) < \infty$, and on the other hand

$$\mathsf{u}(P) - \varrho(P)\,t - \mathsf{s}(P) \geq -\tilde{\mathsf{p}}(t) \quad \text{ for all } t.$$

Inserting the former identity into the last inequality we get the result. As a consequence, the particle density $\varrho(P)$ lies in the interval between the left and right derivative of $\tilde{\mathsf{p}}$ at $\log z$. By convexity, these derivatives are increasing and almost everywhere identical. In fact, they are strictly increasing. For, if they were constant on some non-empty open interval I then for each $t_0 \in I$ and $P \in \mathcal{G}(e^{t_0})$ the function $t \to \varrho(P)\,t - \tilde{\mathsf{p}}(t)$ would be constant on I, and thus by the variational principle $P \in \mathcal{G}(e^t)$ for all $t \in I$. This is impossible because the conditional Gibbs distributions depend non-trivially on the activity. We thus arrive at the following conclusion.

Corollary 4.7 *Let $0 < z < z'$ and $P \in \mathcal{G}(z)$, $P' \in \mathcal{G}(z')$ be translation invariant. Then $\varrho(P) < \varrho(P')$, and $\varrho(P) \to \infty$ as $z \to \infty$.*

In the present two-species model it is natural to consider also the case when each particle species has its own activity, i.e., the plus-particles have activity z^+ and the

minus-particles have activity z^-. It then follows in the same way that the pressure $\mathsf{p}(z^+, z^-)$ is a strictly convex function of the pair $(\log z^+, \log z^-)$, and therefore that the density of plus-particles is a strictly increasing function of z^+, and the density of minus-particles is a strictly increasing function of z^-; these densities tend to infinity as z^+ resp. z^- tends to infinity. Moreover, Theorem 4.3 implies that $\mathsf{p}(z^+, z^-)$ has a kink at (z, z) when z is large enough; this means that the convex function $t \to \mathsf{p}(ze^t, ze^{-t})$ is not differentiable at $t = 0$.

4.3 Projection on plus-particles

As the continuum Ising model is a two-species model, it is natural to ask what kind of system appears if we forget all minus-particles and only retain the plus-particles. The answer is that their distribution is again Gibbsian for a suitable Hamiltonian. This holds also in the case of different activities z^+ and z^- of plus- and minus-particles, which is the natural context here. To check this, take any box $\Delta \subset \mathbf{R}^d$ and let $\Lambda \supset \Delta$ be so large that the distance of Δ from Λ^c exceeds the range R of J. Integrating over X_Λ^- in (16) and conditioning on $X_{\Lambda\setminus\Delta}^+$ one finds that the conditional distribution of X_Δ^+ for given $X_{\Lambda\setminus\Delta}^+$ under $G_\Lambda(\,\cdot\,|\xi_{\Lambda^c})$ does not depend on Λ and ξ_{Λ^c} and has the Gibbsian form

$$\tilde{G}_\Delta(dX_\Delta^+|X_{\Delta^c}^+) = \tilde{Z}_{\Delta|X_{\Delta^c}^+}^{-1} \exp[-z^-\,\tilde{H}_\Delta(X^+)]\,\pi_\Delta^{z^+}(dX_\Delta^+) \tag{24}$$

with the Hamiltonian

$$\tilde{H}_\Delta(X^+) = \int_\Delta \Big(1 - \exp\Big[-\sum_{x\in X^+} J(x-y)\Big]\Big)\,dy\,. \tag{25}$$

Thus, writing $\mathcal{G}(z^+, z^-)$ for the set of all continuum-Ising Gibbs measures on $\Omega = \mathcal{X}^2$ with Hamiltonian (14) and activities z^+ and z^-, and $\tilde{\mathcal{G}}(z^+, z^-)$ for the set of all Gibbs measures on $\mathcal{X}$ with conditional distributions (24), one obtains the following corollary.

Corollary 4.8 *Let $P \in \mathcal{G}(z^+, z^-)$ and $\tilde{P}$ be the distribution of the configuration of plus-particles. Then $\tilde{P} \in \tilde{\mathcal{G}}(z^+, z^-)$. In particular, $\#\tilde{\mathcal{G}}(z, z) > 1$ when z is large enough.*

The last statement follows from Theorem 4.3.

In the Widom–Rowlinson case when $J = \infty\, 1_{\{|\,\cdot\,|\leq 2r\}}$, the relationship between $\mathcal{G}(z^+, z^-)$ and $\tilde{\mathcal{G}}(z^+, z^-)$ has already been observed in the original paper by Widom and Rowlinson (1970). The plus-Hamiltonian (25) then takes the simple form

$$\tilde{H}_\Delta(X^+) = |\Delta \cap \bigcup_{x\in X^+} B_{2r}(x)|$$

that is, $\tilde{H}_\Delta(X^+)$ is the volume in Δ of the Boolean model with radius $2r$ induced by X^+. In this case, the model with distribution $\tilde{G}_\Delta(dX_\Delta^+|X_{\Delta^c}^+)$ was reinvented by Baddeley and van Lieshout (1995). Having the two-dimensional case in mind, they coined the suggestive term *area-interaction process.* From here one can go one step further to Hamiltonians which use not only the volume but also the other Minkowski functionals. This has been initiated by Likos et al. (1995) and Mecke (1996); see also Mecke's contribution to this volume.

One particularly nice feature of the area-interaction process and its generalization (25) is that it seems to be the only known (non-Poisson) model to which Proposition 3.2 can be applied for establishing positive correlations of increasing functions. This attractiveness property makes the model quite attractive for statistical modelling. (By way of contrast, repulsive point systems can be modelled quite easily, for example by a nonnegative pair interaction.)

However, some caution is necessary due to the phase transition when $z^+ = z^- = z$ is large: The typical configurations of $\tilde{G}_\Delta(\,\cdot\,|\,\emptyset\,)$ for a large finite window Δ then can be typical for either phase, $\tilde{P}^+$ or $\tilde{P}^-$, and thus can have different particle densities. Due to finite size effects, this phenomenon already appears when z^+ and z^- are sufficiently close to each other. So, the spatial statistician should be aware of such an instability of observations and should examine whether this is realistic or not in the situation to be modelled.

Finally, one can use Proposition 3.2 to show that the Gibbs measures $\tilde{P} \in \tilde{\mathcal{G}}(z^+, z^-)$ are stochastically increasing in z^+ and decreasing in z^-. In particular, the density of plus-particles for any $P \in \mathcal{G}(z^+, z^-)$ increases when z^+ increases or z^- decreases, as can also be seen using the partial order (21). The monotonicity results in the last paragraph of Section 4.2 thus follow also from stochastic comparison techniques, but Corollary 4.7 cannot be derived in this way.

4.4 Simulation

There are various reasons for performing Monte–Carlo simulations of physical or statistical systems, as discussed in a number of other contributions to this volume. In the present context, the primary reason is to sharpen the intuition on the system's behavior, so that one can see which properties can be expected to hold. This can lead to conjectures which then hopefully can be checked rigorously.

Here we will show briefly how one can obtain simulation pictures of the continuum Ising model. We start with a continuum Gibbs sampler which is suggested by Proposition 4.2; in the Widom–Rowlinson case it has been proposed by Häggström, van Lieshout and Møller (1997).

Consider a fixed window $\Lambda \subset \mathbf{R}^d$ and the Gibbs distribution $G_\Lambda^{\text{free}} = G_\Lambda(\,\cdot\,|\emptyset, \emptyset)$ for the continuum Ising model in Λ with activity z and free (i.e., empty) boundary condition off Λ. We define a random map $F : \mathcal{X}_\Lambda \to \mathcal{X}_\Lambda$ by the following algorithm:

- take an input configuration $X \in \mathcal{X}_\Lambda$,
- select a Poisson configuration $Y \in \mathcal{X}_\Lambda$ with distribution π_Λ^z,
- define a random edge configuration $E \subset \{\{x, y\} : x \in X,\ y \in Y\}$ by independently drawing an edge from $x \in X$ to $y \in Y$ with probability $p(x - y) = 1 - e^{-J(x-y)}$,
- set $F(X) = \Big\{y \in Y : \{x, y\} \notin E \text{ for all } x \in X\Big\}$.

That is, $F(X)$ is a random thinning of π_Λ^z obtained by removing all points which are connected to X by a random edge. Its distribution is nothing other than the

Poisson point process $\pi^{z,J}_{\Lambda|X}$ on Λ with inhomogeneous intensity measure $\rho^J_X(dy) = z\,1_\Lambda(y)\,\exp[-\sum_{x\in X} J(y-x)]\}\,dy$. (Of course, this could also be achieved by setting $F(X) = \{y \in Y : U_y \le \exp[-\sum_{x\in X} J(y-x)]\}$ for independent $U(0,1)$-variables U_y, $y \in Y$. Although this was simpler in the case of the MCMC below, it would considerably increase the running time of the perfect algorithm of Theorem 4.10, as will be explained there.) Now the point is that $\pi^{z,J}_{\Lambda|X^+}$ is the conditional distribution of X^- given X^+ relative to the Gibbs distribution $G^{\rm free}_\Lambda$, and similarly with + and − interchanged. So, if $\xi = (X^+, X^-)$ has distribution $G^{\rm free}_\Lambda$ and F^+, F^- are independent realizations of F then $(F^+(X^-), F^-\circ F^+(X^-))$ has again distribution $G^{\rm free}_\Lambda$. This observation gives rise to the following *Markov chain Monte Carlo algorithm* (MCMC).

Proposition 4.9 *Let F^+_n, F^-_n, $n \ge 0$, be independent realizations of F, and $X^+_0 \in \mathcal{X}_\Lambda$ any initial configuration. Define recursively*

$$X^-_0 = F^-_0(X^+_0), \quad X^+_n = F^+_n(X^-_{n-1}),\ X^-_n = F^-_n(X^+_n) \text{ for } n \ge 1.$$

Then the distribution of (X^+_n, X^-_n) converges to $G^{\rm free}_\Lambda$ in total variation norm at a geometric rate.

Proof. It suffices to observe that $F \equiv \emptyset$ with probability $\delta = e^{-z\,|\Lambda|}$. This shows that for any two configurations $X, X' \in \mathcal{X}_\Lambda$ and any $A \subset \mathcal{X}_\Lambda$

$$|\mathrm{Prob}(F(X)\in A) - \mathrm{Prob}(F(X')\in A)| \le 1-\delta\,.$$

So, if one looks at the process (X^+_n, X^-_n) for two different starting configurations then each application of F reduces the total variation distance by a factor of $1-\delta$. □

A nice property of the random mapping F is its monotonicity: if $X \subset X'$ then $F(X) \supset F(X')$ almost surely. This allows to modify the preceding algorithm to obtain *perfect simulation* in the spirit of Propp and Wilson, as described in the contribution of E. Thönnes to this volume. According to (17), $G^{\rm free}_\Lambda$ is stochastically dominated by independent Poisson processes of plus- and minus-particles. So one can use the idea of *dominated perfect simulation* in her terminology. We describe here only the algorithm and refer to her contribution for more details.

Roughly speaking, the perfect algorithm consists of repeated simultaneous runs of the preceding MCMC, starting from two particular initial conditions at some time $N_k < 0$ until time 0. The two initial conditions are chosen extremal relative to the ordering (21), namely with no initial plus-particle (the minimal case), and with a Poisson crowd of intensity z of plus-particles (which is maximal by stochastic domination). Since the same realizations of F are used in both cases, the two parallel MCMC's have a positive chance of coalescing during the time interval from N_k to 0. If this occurs, one stops. Otherwise one performs a further run which starts at some time $N_{k+1} < N_k$.

Theorem 4.10 *Let F^+_n, F^-_n, $n \le 0$, be independent realizations of F, and $(N_k)_{k\ge1}$ a strictly decreasing sequence of negative run starting times. For each run indexed by $k \ge 1$ let*

$$\Phi_k = F^+_0 \circ F^-_{-1} \circ F^+_{-1} \circ \ldots \circ F^-_{N_k+1} \circ F^+_{N_k+1} \circ F^-_{N_k}$$

be the random mapping corresponding to the MCMC of Proposition 4.9 for the time interval from N_k to 0, and consider the processes

$$X^+_{k,\min} = \Phi_k(\emptyset), \quad X^+_{k,\max} = \Phi_k \circ F^+_{N_k}(\emptyset) \ .$$

Then there exists a smallest (random) $K < \infty$ such that $X^+_{K,\min} = X^+_{K,\max}$, and the random particle configuration $\xi_K = (X^+_{K,\min}, F^-_0(X^+_{K,\min}))$ has distribution G^{free}_Λ.

Since the random mapping F can be simulated by simple standard procedures, the implementation of the preceding algorithm is quite easy; a Macintosh application can be found at `http://www.mathematik.uni-muenchen.de/~georgii/CIsing.html` . The main task is to store the random edge configuration E in each application of F during a time interval $\{N_k, \ldots, N_{k-1} - 1\}$ for use in the later runs (which should be done in a file on the hard disk when $z|\Lambda|$ is large). As a matter of fact, once the set E is determined one can forget the positions of the particles of Y and only keep their indices. In this sense, E contains all essential information of the mapping F. As a consequence, knowing E one needs almost no time to apply the same realization of F in later runs. This is not the case for the alternative definition of F mentioned above.

However, there are some difficulties coming from the phase transition of the model. Running the perfect algorithm for small z is fine and raises no problem. But if z is large then the algorithm requires a considerably longer time to terminate. This is because for each run k the distribution of $\xi_{k,\max} = (X^+_{k,\max}, F^-_0(X^+_{k,\max}))$ will be close to P^+ and thus show a large crowd of plus-particles giving the minus-particles only a minimal chance to spread out. Likewise, the distribution of $\xi_{k,\min} = (X^+_{k,\min}, F^-_0(X^+_{k,\min}))$ will be close to P^-, so that the minus-particles are in the great majority. The bottleneck between these two types of configurations is so small that K will typically be much too large for practical purposes, at least for windows Λ of satisfactory size. In order to reduce this difficulty, one should not simulate the Gibbs distribution G^{free}_Λ with free boundary condition (as we have done above for simplicity), as this distributes most of its mass on two quite opposed events. Rather one should simulate one of the phases, say P^+. For a finite window, this can be achieved by imposing a random boundary condition of Poisson plus-particles outside Λ as in (18). Such a boundary condition helps $X^+_{k,\max}$ and $X^+_{k,\min}$ quite a lot to coalesce within reasonable time. (If one is willing to accept long running times, one should impose periodic boundary conditions to reduce the finite-size effects.)

This electronic version does not contain the figures,
which can be found on the website mentioned above.

The pictures shown are obtained in this way. The underlying interaction potential is $J(x) = 3\,(1 - |x|)^2$ for $|x| \le 1$, $J(x) = 0$ otherwise. The size of the window is 20×20. Outside of the window there is an invisible boundary condition of white Poisson particles. The activities are $z = 4.0$ (top) and $z = 4.5$ (bottom), and the corresponding coalescence times are $-N_K = 300$ resp. 400.

In the subcritical case $z = 4$, the particles in the bulk do not feel the boundary condition: there is no dominance of white over black. In the supercritical case $z = 4.5$

however, the influence of the white boundary condition is strong enough to dominate the whole window, and the phase transition becomes manifest. This is nicely illustrated by the random-cluster representation, which according to Proposition 4.2 is obtained from the point configuration (X^+, X^-) by adding random edges within X^+ and X^- separately. Here one sees that in the subcritical case the influence of the plus-boundary condition is only felt by a the particles near the boundary, while in the supercritical case the global behavior is dominated by a macroscopic cluster reaching from the boundary far into the interior of Λ. This visualizes the equivalence of phase transition and percolation derived in Corollary 4.4.

To conclude we mention two other algorithms. First, there is another perfect algorithm using a rejection scheme due to Fill, which has been studied in detail by Thönnes (1999) in the Widom-Rowlinson case; its extension to the present case is straightforward. A further possibility, which is particularly useful in the supercritical case, is to use a continuum analog of the Swendsen–Wang algorithm, Swendsen and Wang (1987). In its classical version for the lattice Ising model, this algorithm consists in alternating applications of the two procedures in Proposition 2.3. In its continuum version, one can again alternate between the two procedures of Proposition 4.2, but one has to combine this with applications of the random mapping F in order to obtain a resampling of particle positions. Unfortunately, this algorithm does not seem to admit a perfect version because of its lack of monotonicity, but it has the advantage of working also for the many-species extension of the model.

References

Baddeley, A.J. and van Lieshout, M.N.M. (1995) Area-interaction point processes, *Ann. Inst. Statist. Math.* **46**, 601–619.

Chayes, J.T., Chayes, L. and Kotecký, R. (1995) The analysis of the Widom-Rowlinson model by stochastic geometric methods, *Commun. Math. Phys.* **172**, 551–569.

Edwards, R.G. and Sokal, A.D. (1988) Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm, *Phys. Rev.* D **38**, 2009–2012.

Fortuin, C.M. and Kasteleyn, P.W. (1972) On the random-cluster model. I. Introduction and relation to other models, *Physica* **57**, 536–564.

Georgii, H.-O. (1988) *Gibbs Measures and Phase Transitions*, de Gruyter, Berlin New York.

Georgii, H.-O. (1994) Large Deviations and the Equivalence of Ensembles for Gibbsian Particle Systems with Superstable Interaction. *Probab. Th. Rel. Fields* **99**, 171–195.

Georgii, H.-O. (1995) The Equivalence of Ensembles for Classical Systems of Particles. *J. Statist. Phys.* **80**, 1341–1378.

Georgii, H.-O. (1999) Translation invariance and continuous symmetries in two-dimensional continuum systems, in: S. Miracle-Sole, J. Ruiz, V. Zagrebnov (eds.), *Mathematical results in Statistical Mechanics*, Singapore etc., World Scientific, pp. 53–69.

Georgii, H.-O. and Häggström, O. (1996) Phase transition in continuum Potts models, *Commun. Math. Phys.* **181**, 507–528.

Georgii, H.-O., Häggström, O. and Maes, C. (1999) The random geometry of equilibrium phases, in: Domb and J.L. Lebowitz (eds.), *Critical phenomena*, Academic Press, forthcoming.

Georgii, H.-O. and Higuchi, Y. (1999) Percolation and number of phases in the 2D Ising model, submitted to *J. Math. Phys.*

Georgii, H.-O. and Küneth, T. (1997) Stochastic comparison of point random fields, *J. Appl. Probab.* **34**, 868–881.

Grimmett, G.R. (1999) *Percolation*, 2. ed., Springer, New York.

Gruber, Ch. and Griffiths, R.B. (1986) Phase transition in a ferromagnetic fluid, *Physica A* **138**, 220–230.

Häggström, O., van Lieshout, M.N.M. and Møller, J. (1997) Characterization results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes, *Bernoulli*, to appear.

Likos, C.N, Mecke, K.R. and Wagner, H. (1995) Statistical morphology of random interfaces in microemulsions, *J. Chem. Phys.* **102**, 9350–9361.

Mecke, K.R. (1996) A morphological model for complex fluids, *J. Phys.: Condens. Matter* **8**, 9663–9668.

Meester, R. and Roy, R. (1996) *Continuum Percolation*, Cambridge University Press.

Penrose, M.D. (1991): On a continuum percolation model, *Adv. Appl. Probab.* **23**, 536–556.

Swendsen, R.H. and Wang, J.-S. (1987) Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.* **58**, 86–88.

Thönnes, E. (1999): Perfect simulation of some point processes for the impatient user, *Adv. Appl. Probab.* **31**, 69–87.

Widom, B. and Rowlinson, J.S. (1970) New model for the study of liquid-vapor phase transition, *J. Chem. Phys.* **52**, 1670–1684.